

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: WOZNEY, John
CELESTE, Anthony J.
THIES, R. Scott
- (ii) TITLE OF INVENTION: BMP-11 COMPOSITIONS
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: GENETICS INSTITUTE, INC.
 - (B) STREET: 87 CambridgePark Drive
 - (C) CITY: Cambridge
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: LAZAR, Steven R.
 - (B) REGISTRATION NUMBER: 32,618
 - (C) REFERENCE/DOCKET NUMBER: GI5205B-PCT
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 617 498-8260
 - (B) TELEFAX: 617 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 789 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Bos Taurus
 - (B) STRAIN: Bovine Activin WC
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 324..704
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 322..323

(D) OTHER INFORMATION: /note= "putative 3' end of intron"

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 375..701

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAACTGTATT TTGGGGTGAA GGTGTGAGTT AATAGATTCA CGGGACAACA AAGATGGGCT 60
GTTGTTGAGA CCTTGGGCCA AGGGGCTGAT GAGGGTCAGG TTGCCAAGAG AGAGAGAATT 120
AGGGAAGGTG AGTTTAGGGA GACATGGCTA GCTGGCAAGA AAAGTGGGTA GAAAACAGGG 180
GTTGGGGAGG GGAGCACTGG AGAAGCTCAG AAATCACTTG GTCTCTGTTC TCCTGCCCCT 240
ACTGAGGGGC AGGTGAGAAG AAACAGGGAG TAGGAGCTCC TCGAGGCTCT ATTACATCTC 300
TTTCTCCTCT CCCTCACCCC CAG CAT CCT TTT ATG GAG CTT CGA GTC CTA 350
His Pro Phe Met Glu Leu Arg Val Leu
-17 -15 -10
GAG AAC ACA AAA CGG TCC CGG CGG AAC CTG GGC CTG GAC TGC GAT GAA 398
Glu Asn Thr Lys Arg Ser Arg Arg Asn Leu Gly Leu Asp Cys Asp Glu
-5 1 5
CAT TCA AGT GAG TCC CGC TGT TGC CGC TAC CCC CTC ACT GTG GAC TTT 446
His Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro Leu Thr Val Asp Phe
10 15 20
GAG GCT TTT GGC TGG GAC TGG ATC ATC GCT CCT AAA CGC TAC AAG GCC 494
Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro Lys Arg Tyr Lys Ala
25 30 35 40
AAC TAC TGC TCC GGC CAG TGC GAG TAC ATG TTT ATG CAA AAG TAT CCG 542
Asn Tyr Cys Ser Gly Gln Cys Glu Tyr Met Phe Met Gln Lys Tyr Pro
45 50 55
CAC ACC CAC TTG GTG CAA CAG GCT AAC CCA AGA GGC TCT GCG GGG CCC 590
His Thr His Leu Val Gln Gln Ala Asn Pro Arg Gly Ser Ala Gly Pro
60 65 70
TGC TGC ACA CCC ACC AAG ATG TCC CCA ATC AAC ATG CTC TAC TTC AAT 638
Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe Asn
75 80 85
GAC AAG CAG CAG ATT ATC TAC GGC AAG ATC CCT GGC ATG GTG GTG GAT 686
Asp Lys Gln Gln Ile Ile Tyr Gly Lys Ile Pro Gly Met Val Val Asp
90 95 100
CGC TGT GGC TGC TCC TAAGGTGGGG GACAGCGGAT GCCTCCCCAA CAGACCCTGC 741
Arg Cys Gly Cys Ser
105 110
CCCTAGACTC CCCAGCCCT GACCCCTGC TCCCCGGCCC TAGAGCTC 789

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 126 amino acids

(B) TYPE: amino acid

10029016 1200T
T00223T " 9T06200T

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
His Pro Phe Met Glu Leu Arg Val Leu Glu Asn Thr Lys Arg Ser Arg
-17      -15              -10              -5

Arg Asn Leu Gly Leu Asp Cys Asp Glu His Ser Ser Glu Ser Arg Cys
      1              5              10              15

Cys Arg Tyr Pro Leu Thr Val Asp Phe Glu Ala Phe Gly Trp Asp Trp
              20              25              30

Ile Ile Ala Pro Lys Arg Tyr Lys Ala Asn Tyr Cys Ser Gly Gln Cys
              35              40              45

Glu Tyr Met Phe Met Gln Lys Tyr Pro His Thr His Leu Val Gln Gln
              50              55              60

Ala Asn Pro Arg Gly Ser Ala Gly Pro Cys Cys Thr Pro Thr Lys Met
      65              70              75

Ser Pro Ile Asn Met Leu Tyr Phe Asn Asp Lys Gln Gln Ile Ile Tyr
      80              85              90              95

Gly Lys Ile Pro Gly Met Val Val Asp Arg Cys Gly Cys Ser
              100              105
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo Sapiens
- (B) STRAIN: Human Activin WC

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 28..183

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 184..185
- (D) OTHER INFORMATION: /note= "two-thirds of codon at end of partial clone"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
TCTAGATGCT CCGGCCAGTG CGAGTAC ATG TTC ATG CAA AAA TAT CCG CAT
Met Phe Met Gln Lys Tyr Pro His
      1              5
```

51

ACC CAT TTG GTG CAG CAG GCC AAT CCA AGA GGC TCT GCT GGG CCC TGT 99
 Thr His Leu Val Gln Gln Ala Asn Pro Arg Gly Ser Ala Gly Pro Cys
 10 15 20

TGT ACC CCC ACC AAG ATG TCC CCA ATC AAC ATG CTC TAC TTC AAT GAC 147
 Cys Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe Asn Asp
 25 30 35 40

AAG CAG CAG ATT ATC TAC GGC AAG ATC CCT GGC ATG GTGGTGGATC 193
 Lys Gln Gln Ile Ile Tyr Gly Lys Ile Pro Gly Met
 45 50

GCTGTGGCTG CTCCGGATCC 213

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Phe Met Gln Lys Tyr Pro His Thr His Leu Val Gln Gln Ala Asn
 1 5 10 15

Pro Arg Gly Ser Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro
 20 25 30

Ile Asn Met Leu Tyr Phe Asn Asp Lys Gln Gln Ile Ile Tyr Gly Lys
 35 40 45

Ile Pro Gly Met
 50

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: primer C to Bovine Activin WC

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..9
- (D) OTHER INFORMATION: /note= "Restriction site for XbaI"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TAGTCTAGAT GCTCCGGCCA GTGCGAGTAC

30

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Primer D to Bovine Activin WC
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..9
 - (D) OTHER INFORMATION: /note= "Restriction site for BamHI"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGCGGATCCG GAGCAGCCAC AGCGATCCAC

30

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: DNA inserted into pMT2 CXM
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CATGGGCAGC TCGAG

15

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: DNA inserted into pMT21
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..6
 - (D) OTHER INFORMATION: /note= "Pst restriction site"
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 15..26

(D) OTHER INFORMATION: /note= "Eco RI and XhoI restriction sites"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGCAGGCGA GCCTGAATTC CTCGAGCCAT CATG

34

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 68 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Portion of the EMC virus leader sequence

(x) PUBLICATION INFORMATION:

- (A) AUTHORS: Jung, S K
- (C) JOURNAL: J. Virol.
- (D) VOLUME: 63
- (F) PAGES: 1651-1660
- (G) DATE: 1989

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGAGGTTAAA AAACGTCTAG GCGGGGCGAA CCACGGGGAC GTGGTTTTC TTTGAAAAAC

60

ACGATTGC

68

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1270 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human BMP-11

(vii) IMMEDIATE SOURCE:

- (B) CLONE: FB30.5

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1086

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 760..1086

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAG Glu -253	CGC Arg	TCC Ser	AGC Ser -250	CGG Arg	CCA Pro	GCC Ala	CCG Pro	TCC Ser -245	GTG Val	GCG Ala	CCC Pro	GAG Glu	CCG Pro -240	GAC Asp	GGC Gly	48
TGC Cys	CCC Pro	GTG Val -235	TGC Cys	GTT Val	TGG Trp	CGG Arg	CAG Gln -230	CAC His	AGC Ser	CGC Arg	GAG Glu	CTG Leu -225	CGC Arg	CTA Leu	GAG Glu	96
AGC Ser	ATC Ile -220	AAG Lys	TCG Ser	CAG Gln	ATC Ile	TTG Leu -215	AGC Ser	AAA Lys	CTG Leu	CGG Arg	CTC Leu -210	AAG Lys	GAG Glu	GCG Ala	CCC Pro	144
AAC Asn -205	ATC Ile	AGC Ser	CGC Arg	GAG Glu	GTG Val -200	GTG Val	AAG Lys	CAG Gln	CTG Leu	CTG Leu -195	CCC Pro	AAG Lys	GCG Ala	CCG Pro	CCG Pro -190	192
CTG Leu	CAG Gln	CAG Gln	ATC Ile	CTG Leu -185	GAC Asp	CTA Leu	CAC His	GAC Asp	TTC Phe -180	CAG Gln	GGC Gly	GAC Asp	GCG Ala	CTG Leu -175	CAG Gln	240
CCC Pro	GAG Glu	GAC Asp	TTC Phe -170	CTG Leu	GAG Glu	GAG Glu	GAC Asp	GAG Glu -165	TAC Tyr	CAC His	GCC Ala	ACC Thr	ACC Thr -160	GAG Glu	ACC Thr	288
GTC Val	ATT Ile	AGC Ser -155	ATG Met	GCC Ala	CAG Gln	GAG Glu	ACG Thr -150	GAC Asp	CCA Pro	GCA Ala	GTA Val	CAG Gln -145	ACA Thr	GAT Asp	GGC Gly	336
AGC Ser	CCT Pro -140	CTC Leu	TGC Cys	TGC Cys	CAT His	TTT Phe -135	CAC His	TTC Phe	AGC Ser	CCC Pro	AAG Lys -130	GTG Val	ATG Met	TTC Phe	ACA Thr	384
AAG Lys -125	GTA Val	CTG Leu	AAG Lys	GCC Ala	CAG Gln -120	CTG Leu	TGG Trp	GTG Val	TAC Tyr	CTA Leu -115	CGG Arg	CCT Pro	GTA Val	CCC Pro	CGC Arg -110	432
CCA Pro	GCC Ala	ACA Thr	GTC Val	TAC Tyr -105	CTG Leu	CAG Gln	ATC Ile	TTG Leu -100	CGA Arg	CTA Leu	AAA Lys	CCC Pro	CTA Leu	ACT Thr -95	GGG Gly	480
GAA Glu	GGG Gly	ACC Thr	GCA Ala -90	GGG Gly	GGA Gly	GGG Gly	GGC Gly	GGA Gly -85	GGC Gly	CGG Arg	CGT Arg	CAC His	ATC Ile -80	CGT Arg	ATC Ile	528
CGC Arg	TCA Ser	CTG Leu -75	AAG Lys	ATT Ile	GAG Glu	CTG Leu	CAC His -70	TCA Ser	CGC Arg	TCA Ser	GGC Gly	CAT His -65	TGG Trp	CAG Gln	AGC Ser	576
ATC Ile	GAC Asp -60	TTC Phe	AAG Lys	CAA Gln	GTG Val	CTA Leu -55	CAC His	AGC Ser	TGG Trp	TTC Phe	CGC Arg -50	CAG Gln	CCA Pro	CAG Gln	AGC Ser	624
AAC Asn -45	TGG Trp	GGC Gly	ATC Ile	GAG Glu	ATC Ile -40	AAC Asn	GCC Ala	TTT Phe	GAT Asp	CCC Pro -35	AGT Ser	GGC Gly	ACA Thr	GAC Asp	CTG Leu -30	672
GCT Ala	GTC Val	ACC Thr	TCC Ser	CTG Leu -25	GGG Gly	CCG Pro	GGA Gly	GCC Ala	GAG Glu -20	GGG Gly	CTG Leu	CAT His	CCA Pro	TTC Phe -15	ATG Met	720
GAG Glu	CTT Leu	CGA Arg	GTC Val -10	CTA Leu	GAG Glu	AAC Asn	ACA Thr	AAA Lys -5	CGT Arg	TCC Ser	CGG Arg	CGG Arg	AAC Asn	CTG Leu	GGT Gly	768

CTG GAC TGC GAC GAG CAC TCA AGC GAG TCC CGC TGC TGC CGA TAT CCC	816
Leu Asp Cys Asp Glu His Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro	
5 10 15	
CTC ACA GTG GAC TTT GAG GCT TTC GGC TGG GAC TGG ATC ATC GCA CCT	864
Leu Thr Val Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro	
20 25 30 35	
AAG CGC TAC AAG GCC AAC TAC TGC TCC GGC CAG TGC GAG TAC ATG TTC	912
Lys Arg Tyr Lys Ala Asn Tyr Cys Ser Gly Gln Cys Glu Tyr Met Phe	
40 45 50	
ATG CAA AAA TAT CCG CAT ACC CAT TTG GTG CAG CAG GCC AAT CCA AGA	960
Met Gln Lys Tyr Pro His Thr His Leu Val Gln Gln Ala Asn Pro Arg	
55 60 65	
GGC TCT GCT GGG CCC TGT TGT ACC CCC ACC AAG ATG TCC CCA ATC AAC	1008
Gly Ser Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn	
70 75 80	
ATG CTC TAC TTC AAT GAC AAG CAG CAG ATT ATC TAC GGC AAG ATC CCT	1056
Met Leu Tyr Phe Asn Asp Lys Gln Gln Ile Ile Tyr Gly Lys Ile Pro	
85 90 95	
GGC ATG GTG GTG GAT CGC TGT GGC TGC TCT TAAGGTGGGG GATAGAGGAT	1106
Gly Met Val Val Asp Arg Cys Gly Cys Ser	
100 105	
GCCTCCCCCA CAGACCCTAC CCCAAGACCC CTAGCCCTGC CCCCATCCCC CCAAGCCCTA	1166
GAGCTCCCTC CACTCTTCCC GCGAACATCA CACCGTTCCC CGACCAAGCC GTGTGCAATA	1226
CAACAGAGGG AGGCAGGTGG GAATTGAGGG TGAGGGGTTT GGGG	1270

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 362 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Glu Arg Ser Ser Arg Pro Ala Pro Ser Val Ala Pro Glu Pro Asp Gly	
-253 -250 -245 -240	
Cys Pro Val Cys Val Trp Arg Gln His Ser Arg Glu Leu Arg Leu Glu	
-235 -230 -225	
Ser Ile Lys Ser Gln Ile Leu Ser Lys Leu Arg Leu Lys Glu Ala Pro	
-220 -215 -210	
Asn Ile Ser Arg Glu Val Val Lys Gln Leu Leu Pro Lys Ala Pro Pro	
-205 -200 -195 -190	
Leu Gln Gln Ile Leu Asp Leu His Asp Phe Gln Gly Asp Ala Leu Gln	
-185 -180 -175	

Pro Glu Asp Phe Leu Glu Glu Asp Glu Tyr His Ala Thr Thr Glu Thr
-170 -165 -160

Val Ile Ser Met Ala Gln Glu Thr Asp Pro Ala Val Gln Thr Asp Gly
-155 -150 -145

Ser Pro Leu Cys Cys His Phe His Phe Ser Pro Lys Val Met Phe Thr
-140 -135 -130

Lys Val Leu Lys Ala Gln Leu Trp Val Tyr Leu Arg Pro Val Pro Arg
-125 -120 -115 -110

Pro Ala Thr Val Tyr Leu Gln Ile Leu Arg Leu Lys Pro Leu Thr Gly
-105 -100 -95

Glu Gly Thr Ala Gly Gly Gly Gly Gly Gly Arg Arg His Ile Arg Ile
-90 -85 -80

Arg Ser Leu Lys Ile Glu Leu His Ser Arg Ser Gly His Trp Gln Ser
-75 -70 -65

Ile Asp Phe Lys Gln Val Leu His Ser Trp Phe Arg Gln Pro Gln Ser
-60 -55 -50

Asn Trp Gly Ile Glu Ile Asn Ala Phe Asp Pro Ser Gly Thr Asp Leu
-45 -40 -35 -30

Ala Val Thr Ser Leu Gly Pro Gly Ala Glu Gly Leu His Pro Phe Met
-25 -20 -15

Glu Leu Arg Val Leu Glu Asn Thr Lys Arg Ser Arg Arg Asn Leu Gly
-10 -5 1

Leu Asp Cys Asp Glu His Ser Ser Glu Ser Arg Cys Cys Arg Tyr Pro
5 10 15

Leu Thr Val Asp Phe Glu Ala Phe Gly Trp Asp Trp Ile Ile Ala Pro
20 25 30 35

Lys Arg Tyr Lys Ala Asn Tyr Cys Ser Gly Gln Cys Glu Tyr Met Phe
40 45 50

Met Gln Lys Tyr Pro His Thr His Leu Val Gln Gln Ala Asn Pro Arg
55 60 65

Gly Ser Ala Gly Pro Cys Cys Thr Pro Thr Lys Met Ser Pro Ile Asn
70 75 80

Met Leu Tyr Phe Asn Asp Lys Gln Gln Ile Ile Tyr Gly Lys Ile Pro
85 90 95

Gly Met Val Val Asp Arg Cys Gly Cys Ser
100 105